Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **470,** 1809–1825 (2017)
Advance Access publication 2017 May 29

doi:10.1093/mnras/stx1330

Hubble Frontier Fields: systematic errors in strong lensing models of galaxy clusters – implications for cosmography

Ana Acebron,[1]★ Eric Jullo,[1] Marceau Limousin,[1] André Tilquin,[2] Carlo Giocoli,[1,3] Mathilde Jauzac,[4,5,6] Guillaume Mahler[7] and Johan Richard[7]

[1]*Aix Marseille Université, CNRS, LAM, Laboratoire d'Astrophysique de Marseille, Marseille, France*
[2]*Aix Marseille Université, CNRS, CPPM, Marseille, F-13288 Marseille cedex 09, France*
[3]*Dipartimento di Fisica e Astronomia, Alma Mater Studiorum Università di Bologna, via Gobetti 93/2, I-40129 Bologna, Italy*
[4]*Centre for Extragalactic Astronomy, Department of Physics, Durham University, Durham DH1 3LE, UK*
[5]*Institute for Computational Cosmology, Departement of Physics, University of Durham, South Road, Durham DH1 3LE, UK*
[6]*Astrophysics and Cosmology Research Unit, School of Mathematical Sciences, University of KwaZulu-Natal, Durban 4041, South Africa*
[7]*Université Lyon, Université Lyon1, Ens de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, F-69230 Saint-Genis-Laval, France*

Accepted 2017 May 25. Received 2017 May 23; in original form 2017 April 18

ABSTRACT

Strong gravitational lensing by galaxy clusters is a fundamental tool to study dark matter and constrain the geometry of the Universe. Recently, the *Hubble Space Telescope* Frontier Fields programme has allowed a significant improvement of mass and magnification measurements but lensing models still have a residual root mean square between 0.2 arcsec and few arcseconds, not yet completely understood. Systematic errors have to be better understood and treated in order to use strong lensing clusters as reliable cosmological probes. We have analysed two simulated *Hubble-Frontier-Fields*-like clusters from the *Hubble Frontier Fields* Comparison Challenge, *Ares* and *Hera*. We use several estimators (relative bias on magnification, density profiles, ellipticity and orientation) to quantify the goodness of our reconstructions by comparing our multiple models, optimized with the parametric software LENSTOOL, with the input models. We have quantified the impact of systematic errors arising, first, from the choice of different density profiles and configurations and, secondly, from the availability of constraints (spectroscopic or photometric redshifts, redshift ranges of the background sources) in the parametric modelling of strong lensing galaxy clusters and therefore on the retrieval of cosmological parameters. We find that substructures in the outskirts have a significant impact on the position of the multiple images, yielding tighter cosmological contours. The need for wide-field imaging around massive clusters is thus reinforced. We show that competitive cosmological constraints can be obtained also with complex multimodal clusters and that photometric redshifts improve the constraints on cosmological parameters when considering a narrow range of (spectroscopic) redshifts for the sources.

Key words: gravitational lensing: strong – galaxies: clusters: general – cosmological parameters.

1 INTRODUCTION

Strong gravitational lensing (SL, hereafter) is nowadays at the very heart of important issues in modern cosmology, allowing us, for instance, to measure directly the total projected mass distribution (baryonic and dark) of galaxy cluster cores (e.g. Richard et al. 2010; Zitrin & Broadhurst 2016; Mahler et al. 2017; Monna et al. 2017), to image very high redshift sources otherwise too faint to be detected without the gravitational magnification (Kneib et al. 2004; Richard et al. 2008; Coe et al. 2013; Atek et al. 2015) and to constrain the geometry of the Universe (Jullo et al. 2010; D'Aloisio & Natarajan 2011; Magaña et al. 2015; Caminha et al. 2016), which is one of the long-standing challenges of modern cosmology.

In the standard cosmological model Lambda cold dark matter (ΛCDM), $\sim$72 per cent of the energy density of the Universe is in the form of a 'dark energy', a fluid with negative pressure that would cause the presently acceleration of the Universe (Planck Collaboration XIII 2016a).

Type Ia supernovae (Riess et al. 1998), baryon acoustic oscillations (Beutler et al. 2011; Gil-Marín et al. 2016), cosmic shear (Massey et al. 2007; Heymans et al. 2012; Hildebrandt et al. 2017),

★ E-mail: ana.acebron@lam.fr

cluster abundances (de Haan et al. 2016), cosmic microwave background anisotropies (Planck Collaboration XIII 2016a) or time delays (Suyu et al. 2017) are some of the several probes allowing us to better understand the constituents of the Universe and their properties by putting tighter constraints on cosmological parameters.

However, in order to obtain robust estimates of cosmological parameters, estimates from the different cosmological probes must be combined as each technique has distinct degeneracies and biases (Planck Collaboration XIV 2016b; Peel et al. 2017).

Among the previously mentioned cosmological probes, using the strong lensing features in galaxy clusters is a very promising technique that yields orthogonal constraints in an era of precise cosmology (Golse, Kneib & Soucail 2002; Gilmore & Natarajan 2009; Jullo et al. 2010). To perform cosmography with strong lenses, a precise and accurate mass distribution of the cluster is required, i.e. a large number of constraints is crucial.

Recently, the *Hubble Frontier Fields* programme[1] (HFF; P.I.: J. Lotz; Lotz et al. 2017) with the *Hubble Space Telescope* (hereafter *HST*) has provided the deepest multicolour imaging of galaxy clusters to date, which, combined with spectroscopy from ground-based surveys, has led to the discovery of hundreds of multiple images and thus to a significant improvement of cluster mass estimates (Jauzac et al. 2014; Diego et al. 2016; Lagattuta et al. 2017; Monna et al. 2017).

The mass modelling of strong lensing clusters can be carried out in different manners: Parametric and non-parametric methods are equally used; the primary distinction between them being that parametric modelling assumes that luminous cluster galaxies trace the cluster mass whereas non-parametric does not.

The FF-SIMS Challenge (Meneghetti et al. 2016), an archive of mock HFF-like clusters, has provided the lensing community with a set of simulated clusters in order to highlight for the first time the strengths and weaknesses of each methods: parametric and non-parametric. This work would allow the non-lensing community to choose a method and software according to their different needs. This challenge has shown that all lensing reconstruction methods provide reliable mass distributions. However, strong lensing modelling appears to be still unable to match the *HST* observation angular resolution ($\sim$0.05 arcsec) with a residual root mean square (rms) between 0.2 arcsec and a few arcseconds (Limousin et al. 2016), a systematical error not yet completely understood, and few studies have addressed this issue.

Indeed, strong lensing mass modelling has various sources of systematic errors, arising from the hypothesis behind our models, which have recently started to be acknowledged and analysed in a more quantitative way. Meneghetti et al. (2010) studied the properties of $\sim$50 000 strong lensing clusters in the MARENOSTRUM cosmological simulation. They find that strong lenses tend to have their major axes oriented along the line of sight. This orientation bias results, for instance, in cluster concentrations estimations from the projected density profiles to be biased high (Giocoli et al. 2014; Sereno et al. 2015). D'Aloisio & Natarajan (2011) quantified that the modelling errors due to the scatter in the cluster–galaxy scaling relations and unmodelled line-of-sight haloes can result in errors of the order of a few arcseconds on average. Bayliss, Sharon & Johnson (2015) studied the impact of assuming a certain cosmological model on the determination of magnification and mass profiles of the cluster core for HFF clusters showing that cosmological parameter uncertainty is a non-negligible source of errors for the lens modelling.

Foreground or background large-scale structures impacts the SL modelling as well (Dalal, Hennawi & Bode 2005; Host 2012), which could introduce a systematic error of up to $\sim$1.4 arcsec on the position of multiple images (Zitrin et al. 2015). Line-of-sight effects have then to be taken into account during the strong lensing modelling in order to recover precise cosmological parameters (Jullo et al. 2010; Caminha et al. 2016).

Systematic errors can also arise from the mass distributions assumed for the DM components as shown in Limousin et al. (2016), where the observed constraints in MACS0717 are equally well reproduced by a mass model with a shallow large-scale DM component and one for which this component is peaky. Harvey, Kneib & Jauzac (2016) analysed the FF cluster MACSJ0416 ($z = 0.397$) and found the assumption that light traces mass can introduce an error of $\sim$0.5 arcsec on the position of the multiple images. Bouwens et al. (2016) studied the impact of magnification uncertainties on luminosity functions from the first four HFF clusters. Johnson & Sharon (2016) have led the first investigation attempting to quantify systematic errors induced by the availability of constraints in strong lensing clusters. They show that the accuracy of the magnification is sensitive to the selection of constraints rather than their amount.

In this paper, we take advantage of the mock HFF-like cluster archive from the FF-SIMS Challenge and use two of the HFF-like mock strong lensing clusters (*Ares* and *Hera*) to investigate how systematic errors in the strong lensing parametric modelling with LENSTOOL affect the determination of the total mass distribution in clusters and hence the retrieval of robust cosmological parameters, such as the mean matter density and dark energy equation-of-state parameters $\Omega_{\rm M}$ and w, respectively.

First, we use four different estimators (density profiles, relative bias on magnification, cluster's ellipticity and orientation angle) to compare our reconstructions with the input models in order to assess the impact of different mass distributions and configurations in the strong lensing modelling of clusters and hence on the cosmography.

Secondly, we study how the availability of constraints affects the retrieval of unbiased cosmological parameters. We investigate if there is a more efficient range of redshifts to recover the input cosmology, if including photometric redshifts can result in a more robust estimation of these parameters and if taking into account an increasing number of photometric families translates into an increasing precision in the cosmological parameters estimation.

This paper is organized as follows: The data used for this analysis are presented in Section 2; in Section 3, we describe the methodology; in Section 4, we detail the different modellings for *Ares* and *Hera*; we then show the systematic uncertainties in Section 5, their impact on cosmography in Section 6 and conclude in Section 7.

Throughout this paper, we use the standard ΛCDM flat cosmological model with the Hubble constant $H_0 = 70\,{\rm km\,s^{-1}\,Mpc^{-1}}$, and $\Omega_{\rm M}$ and w are used as free parameters. Magnitudes are quoted in the AB system.

2 THE DATA

We have analysed two mock HFF-like clusters from the FF-simulations Challenge archive (Meneghetti et al. 2016): *Ares* and *Hera*. They were initially created for the lens modelling comparison project, which, for the first time, compares the reconstructions obtained using different techniques, such as parametric, non-parametric and hybrid, performed by different teams.

[1] http://www.stsci.edu/hst/campaigns/frontier-fields

Both clusters have been simulated to reproduce, not only the characteristics of the *HST* Advandced Camera for Survey (ACS) and Wide-Field Camera 3 (WFC3) observations (depth, passbands and spatial resolution), but also the complexity of the FF clusters themselves even if these two have been created using different techniques.

Both clusters have then been modelled to be realistic, complex, bimodal clusters (see Fig. 1), generated in a flat ΛCDM cosmological model.

During the challenge, all participants first performed a blind analysis (i.e. not knowing the true mass distribution) of these clusters for which only the HFF-like images as well as catalogues of multiple images (position and spectroscopic redshifts) and cluster galaxies (positions and magnitudes in all ACS/WFC3 bands) were provided with $m_{F814W} < 24$ (see Table 1 for further details).

Ares is a more powerful lens than *Hera*, thus producing many more multiple images (Meneghetti et al. 2016). Their respective redshift distributions are shown in Fig. A1. The source galaxies resemble the luminosity and the redshift distribution of the galaxies in the *Hubble Ultra Deep Field* (Coe et al. 2006).

For both clusters, the software SKYLENS (Meneghetti et al. 2008) was used to ray-trace the lensed galaxies to the image plane and to create simulated HFF-like images in all bands.

After the unblinding, the convergence and shear maps (calculated for a source at $z_s = 9$) were provided. They were used to compare our reconstructions with the true values and also to improve our models. The input cosmology was also revealed. We refer the reader to Meneghetti et al. (2016) for a detailed presentation on how *Ares* and *Hera* mass distributions were generated; here, we present only a quick overview.

2.1 *Ares*

Ares is a semi-analytical cluster (created using MOKA[2] by Giocoli et al. 2012) at $z = 0.5$. This simulated cluster is built with three components: two smooth DM triaxial haloes with an NFW profile, a bright central cluster galaxy (BCG) with an Hernquist profile (Hernquist 1990), and sub-haloes having a singular isothermal sphere profile (Hernquist 1990). DM sub-haloes are populated using a halo occupation distribution technique, and stellar and *B*-band luminosities are given for all galaxies according to the mass of their sub-halo as in Wang et al. (2006). The mass within the virial radius is then defined as $M_{\rm vir} = M_{\rm smooth} + \Sigma_i m_{{\rm sub},i}$.

Ares is generated in a flat ΛCDM cosmological model with a matter density parameter $\Omega_{\rm M} = 0.272$.

2.2 *Hera*

Hera is from an *N*-body simulation of cluster-sized DM haloes for a flat ΛCDM model (see Planelles et al. 2014) at $z = 0.507$ that was re-simulated using a TreePM-SPH GADGET-3 code with only collisionless DM particles. The properties of cluster galaxies are created from semi-analytic Methods of galaxy formation (De Lucia & Blaizot 2007). *Hera* is also created in a flat ΛCDM cosmological model with a matter density parameter $\Omega_{\rm M} = 0.24$.

For both clusters, the software SKYLENS (Meneghetti et al. 2008) was used to ray-trace the lensed galaxies to the image plane and to create simulated HFF-like images in all bands.

[2] https://cgiocoli.wordpress.com/research-interests/moka/

3 METHODOLOGY

3.1 Mass modelling

We perform the strong lensing modelling in the source plane using the public software LENSTOOL[3] (Kneib et al. 1996; Jullo et al. 2007), which performed very well for the FF-SIMS Challlenge. LENSTOOL utilizes a Bayesian Markov chain Monte Carlo (hereafter MCMC, see Jullo et al. 2007, for details) sampler to optimize the model using the positions of the multiply imaged systems. The matter distribution of clusters is decomposed into several smooth large-scale components and individual contributions from cluster galaxies. In this work, the reconstructions are based on the strong lensing information only.

We have modelled both clusters as up to three components: (i) large-scale potentials, (ii) BCGs and (iii) individual galaxies identified spectroscopically, with masses scaling with luminosity. Each mass component has a parametrized profile such as the pseudo-isothermal elliptical mass distribution (Kassiola & Kovner 1993, hereafter PIEMD), Navarro, Frenk & White (NFW; Navarro, Frenk & White 1997) or Hernsquist (Hernquist 1990). For instance, large-scale potentials are modelled with either an NFW or PIEMD, BCGs with PIEMD or Hernquist profiles and individual clusters members with a PIEMD profile, scaled according to the relations in Limousin, Kneib & Natarajan (2005). These mass distributions are briefly presented hereafter.

We used the NFW density profile (Navarro et al. 1997) to model *Ares* and *Hera* large-scale mass distributions. The 3D density profile is given by

$$\rho_{\rm NFW}(r) = \frac{\rho_s}{\left(\frac{r}{r_{\rm s}}\right)\left(1+\frac{r}{r_{\rm s}}\right)^2}, \quad (1)$$

where $\rho_{\rm s}$ is a characteristic density and $r_{\rm s}$ is the scale radius.

This profile behaves as $\rho \propto r^{-1}$ in the inner region, $\rho \propto r^{-2}$ at $r = r_{\rm s}$ and as $\rho \propto r^{-3}$ in the outer regions.

In LENSTOOL, this profile has the following free parameters: x and y, the coordinates of the halo centre; e, the ellipticity, defined as $e = (a^2 + b^2)/(a^2 - b^2)$, with a and b being the semimajor and semiminor axis, respectively, and θ, the position angle (counted counter-clockwise from the x-axis); $r_{\rm s}$, the scale radius; and σ, the velocity dispersion.

The pseudo-isothermal density profile (Limousin et al. 2005) is used to model DM haloes and/or individual galaxies where 3D density distribution is given by

$$\rho_{\rm PIEMD}(r) = \frac{\rho_0}{\left(1+\frac{r^2}{{r_{\rm core}}^2}\right)\left(1+\frac{r^2}{{r_{\rm cut}}^2}\right)}, \quad (2)$$

with a core radius $r_{\rm core}$ and a truncation radius $r_{\rm cut}$.

This profile is characterized by two changes in the density slope: within the transition region($r_{\rm core} < r < r_{\rm cut}$), it behaves as an isothermal profile with $\rho \propto r^{-2}$, while exiting this region, the density will fall as $\rho \propto r^{-4}$ (such behaviour is common for elliptical galaxies).

It has the following free parameters: the coordinates x, y; the ellipticity, e; angle position, θ; core and cut radii, $r_{\rm core}$ and $r_{\rm cut}$; and a velocity dispersion, σ.

We use the Hernquist profile (Hernquist 1990) to model the BCGs of the two clusters. BCGs are massive elliptical galaxies with observed luminosities well represented by the de Vaucouleurs $R^{\frac{1}{4}}$ empirical law (which fits well observations) but being more simple

[3] https://projets.lam.fr/projects/lenstool

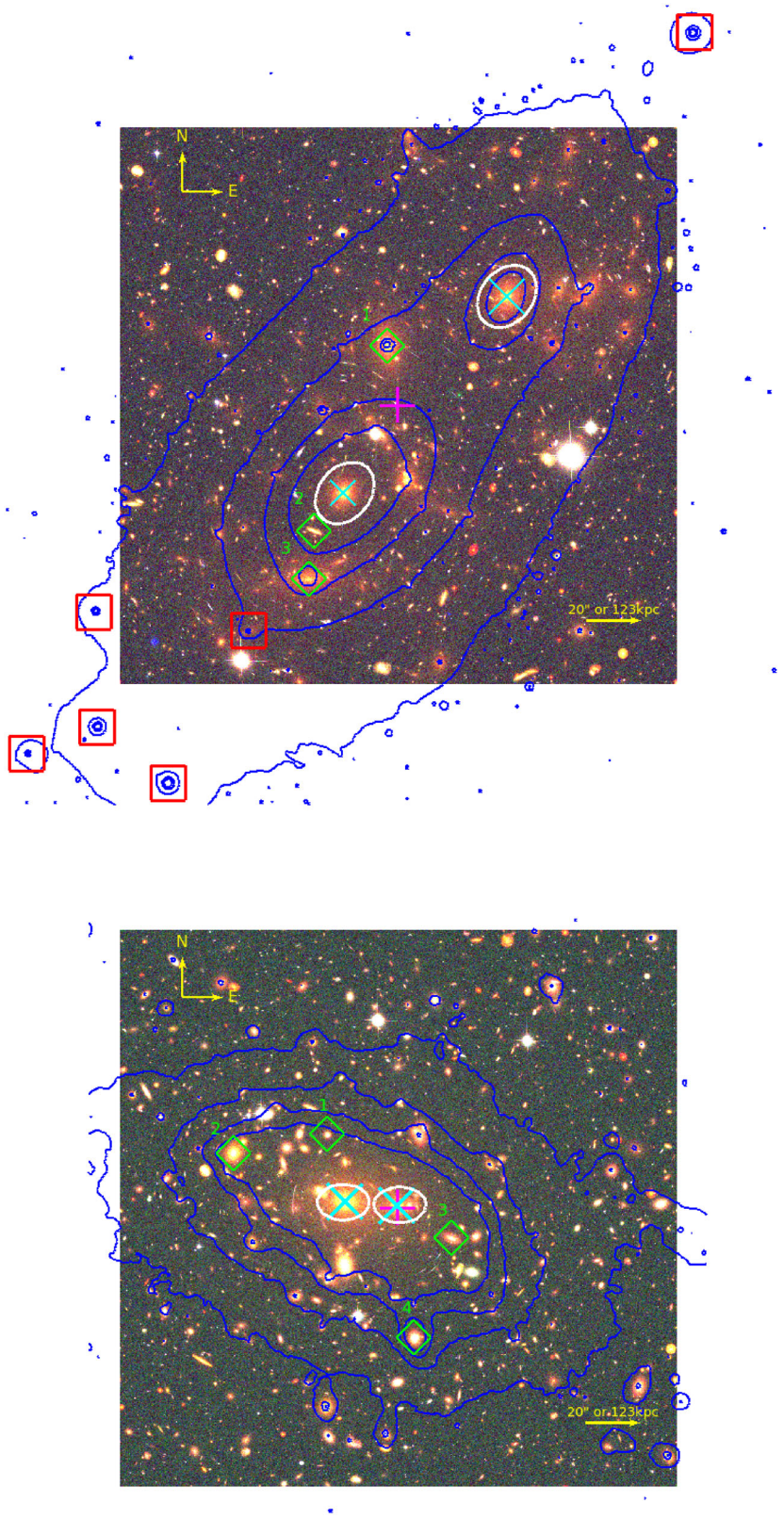

Figure 1. The underlying colour image is a composite created from the provided *HST*/ACS simulated images in the *F814W*, *F606W* and *F435W* passbands. Input convergence maps are shown in blue. Upper panel: *Ares* cluster. The cluster's centre is indicated by the magenta cross. Large-scale potentials are shown as white ellipses, BCGs are shown as cyan crosses, galaxy haloes are shown as green diamonds and red boxes are for the outskirts substructures. Bottom panel: *Hera* cluster. Same colour code as for *Ares*.

analytically. The 3D density profile of the Hernquist profile can be written as follows:

$$\rho_H(r) = \frac{M}{2\pi}\frac{a}{r}\frac{1}{(r+a)^3}, \quad (3)$$

where M is the total mass and a is the characteristic scalelength. This profile is extremely similar to the NFW profile at small radii but the density falls as $\rho \propto r^{-4}$ at larger radii.

It is parametrized using the following free parameters: the coordinates x, y; the ellipticity, e; angle position, θ; a core radius, $r_{\rm core}$; and a velocity dispersion, σ.

Once the mass components are defined, the best-fitting model parameters are found by minimizing the distance between the observed and model-predicted positions of the multiple images, and the parameter covariance is estimated using a Bayesian MCMC technique (Jullo et al. 2007). For each of the models, we use several statistical values to assess the goodness of the fit and to discriminate between models.

(i) We use the rms between the observed and predicted positions of the multiple images from the modelling, computed as follows:

$$\mathrm{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{n} |\theta_i^{\rm obs} - \theta_i^{\rm pred}|^2}, \quad (4)$$

where $\theta_i^{\rm obs}$ and $\theta_i^{\rm pred}$ are the observed and model-predicted positions of the multiples images, with N being the total number of images.

(ii) We compute the Bayesian Information criterion (BIC, introduced by Schwarz 1978):

$$\mathrm{BIC} = -2\ln(L) + k \times \ln(n), \quad (5)$$

with L being the likelihood, k is the number of free parameters and n is the number of constraints.

(iii) The corrected Akaike information criterion (AICc): The BIC will overpenalize models, whereas AICc tends to underpenalize. Using the two together helps balancing those effects. It is computed as follows:

$$\mathrm{AICc} = 2k - \ln(L) + \frac{2k(k+1)}{(n-k-1)}. \quad (6)$$

(iv) The reduced χ^2; see Jullo et al. (2007) for details.

(v) The Bayesian Evidence that considers the 'complexity' of the models and how this 'complexity' is justified by the observables; see also Jullo et al. (2007) for details.

3.2 Cosmological parameters

As both *Ares* and *Hera* have a large number of multiple images with spectroscopic redshifts for all images – up to redshift $z_{\rm s} \sim 6$ for *Ares* and $z_{\rm s} \sim 3.5$ for *Hera* (see Fig. A1) – they represent good probes to constrain cosmological parameters.

We briefly outline here the methodology to estimate these parameters with cluster strong lensing (further details can be found in Golse et al. 2002; Gilmore & Natarajan 2009, for instance).

Strong lensing is sensitive to the underlying geometry of the Universe as the position of the multiple images not only depends on the mass distribution of the lens but also on the angular diameter distances from the observer to the lens ($D_{\rm OL}$), to the source ($D_{\rm OS}$) and from the lens to the source ($D_{\rm LS}$). This dependence is used to put constraints on the cosmological parameters $\Omega_{\rm M}$ and w.

Indeed, the lens equation can be written as

$$\boldsymbol{\beta}_i = \boldsymbol{\theta}_i - \frac{2}{c^2}\frac{D_{\rm LS}}{D_{\rm OL}D_{\rm OS}}\nabla\phi(\boldsymbol{\theta}_i), \quad (7)$$

where θ and β are the (multiple) image angular positions in the lens and source planes, respectively, ϕ is the projected Newtonian potential of the lens and the cosmological dependence is embedded into the angular diameter distances.

When only one family of multiple images is available, the ratio between the cosmological distances cannot be disentangled from the gradient of the potential. However, if at least two systems of images at different redshifts are available, this degeneracy can be broken and via the 'family ratio' (see Link & Pierce 1998) from which constraints on $\Omega_{\rm M}$ and w can be obtained:

$$\Xi(z_{\rm L}, z_{\rm s1}, z_{\rm s2}; \Omega_{\rm M}, w) = \frac{D_{\rm LS_1} D_{\rm OS_2}}{D_{\rm OS_1} D_{\rm LS_2}}, \quad (8)$$

with $z_{\rm L}$ being the redshift of the lens, $z_{\rm s1}$ and $z_{\rm s2}$ are the redshifts of two distinct sources and D is the angular diameter distance.

This kind of analysis has already been carried out for galaxy clusters Abell 2218, Abell 1689 and Abell S1063 (or RXC J2248.7-4431) by Soucail, Kneib & Golse (2004), Jullo et al. (2010) and Caminha et al. (2016), respectively, and also with simulated data (Golse et al. 2002; Gilmore & Natarajan 2009; D'Aloisio & Natarajan 2011).

In this work, the energy density of the total matter of the universe $\Omega_{\rm M}$ and the equation-of-state parameter w are set as free parameters in a ΛCDM flat cosmological model.

3.3 Priors

In this section, we describe the choice of boundaries for the flat priors of the free parameters mentioned below. Some boundaries are set with a quite large interval. However, others have more narrow priors.

The positions of the haloes (for both cluster- and galaxy-scale potentials), x and y, correspond to the light peak of galaxies (the position of the BCGs being the central coordinates for the DM haloes). The angle positions of haloes are also set by the luminous component's angle. The ellipticity is not allowed to reach very high values following Despali et al. (2017). This work shows that high ellipticities ($e > 0.75$) are not favoured by theoretical predictions.

Finally, some parameters have quite narrow priors as models have been run several times, and in order to gain in computing time, some prior ranges were tightened, checking that the boundaries were not reached. None the less, we checked that a slight change in the boundaries had no impact on the posterior distribution of the parameters.

4 STRONG LENSING MODELS

We here present the details of each model for the *Ares* and *Hera* clusters, the free parameters for each potential and the respective flat priors. These models are presented in chronological order as we tried to model the clusters in a more complex way (and also after the unblinding of the true mass profiles).

Their id number throughout this paper is the last number of their corresponding section. The first term stands for the density profile used for the large-scale haloes, the second (if any) stands for BCGs.

All coordinates are presented as the distance to the centre of the clusters (in arcseconds) shown as the magenta crosses in Fig. 1. The centre has been arbitrarily chosen to be (δRA, δDec.) = (0, 0 arcsec) and serves as reference for LENSTOOL.

4.1 *Ares*

Thanks to the *HST*-simulated images, we easily see that this cluster is bimodal; all the models thereafter will have two large-scale haloes whose coordinates stated below correspond to those of the light peaks (see the upper panel of Fig. 1). Cluster member galaxies are taken from the given simulated catalogues up to a magnitude of $m_{F160W} < 22.0$ (being a more complex cluster, we limited the sample with a magnitude cut to gain in computing time, representing >90 per cent of the total cluster luminosity) and the small-scale haloes associated with galaxy members are parametrized with a PIEMD profile with a fixed core radius of 0.15 kpc, a velocity dispersion $\sigma^\star$ allowed to vary between 94 and 180 km s^{-1}, a cut radius $r^\star_{\rm cut}$ varying from 78 to 272.00 kpc, considered spherical, with a $mag0 = 18.5$ (the reference magnitude for the scaling relations corresponding to $L^\star$ at the cluster's redshift) and following the scaling relations (Faber & Jackson 1976).

Moreover, three massive galaxies were more carefully modelled (shown as green diamonds in Fig. 1) also using a PIEMD density profile. These massive galactic haloes were modelled independently (not using the scaling relations) as they appeared to have a significant impact in the mass profile and/or on the nearby multiple images. For the three of them, the cut radius $r^\star_{\rm cut}$ is fixed to 1000 kpc, as initially the modelling was undertaken as the merger of four haloes (supported by the high velocity dispersions of these massive galaxies).

The first galaxy (labelled as 1 in Fig. 1) is fixed at (δRA, δDec.) = (+4.053 arcsec, +22.042 arcsec) away from the centre of the cluster. Its ellipticity can go up to 0.6; the angle position is allowed to vary from 60.0 to 120.0 and its velocity dispersion σ is allowed to vary from 100.0 to 400.0 km s^{-1}.

The galaxy labelled as 2 in Fig. 1 is located at (δRA, δDec.) = (+30.78 arcsec, −45.98 arcsec) away from the centre of the cluster with an ellipticity from 0.3 to 0.7; the angle position is allowed to vary from 0.0 to 90.0 and its velocity dispersion σ is allowed to vary from 50.0 to 400.0 km s^{-1}.

Finally, the galaxy labelled as 3 in Fig. 1 is located at (δRA, δDec.) = (+33.008 arcsec, −63.542 arcsec) away from the centre of the cluster. From the input convergence contours, it appeared as a massive halo. It is modelled with ellipticity with values between 0.1 and 0.7; the angle position is allowed to vary from 141° to 171°, and its velocity dispersion σ is allowed to vary from 400.0 to 700.0 km s^{-1}.

All multiple images are included in the models with a positional uncertainty of 0.5 arcsec and the optimization is performed in the source plane, as it is less computing time expensive. We checked that the results are similar with both source and image plane optimizations.

4.1.1 *PIEMD*

This first model contains two large-scale haloes whose coordinates are fixed at (δRA, δDec.) = (+20.0 arcsec, −32.0 arcsec) and (−40.0 arcsec, +40.0 arcsec) away from the centre of the cluster, respectively. These clumps do not harbour any additional halo linked to the BCGs.

For the first clump, we let the core radius vary from 20 to 65 kpc with a fixed cut radius of 1000 kpc. The ellipticity of this halo is allowed to reach values from 0.2 and as high as 0.7, its orientation can vary from 130° to 140° and the velocity dispersion can vary from 400 to 1700 km s^{-1}. The second large-scale halo can have a core radius with any value from 20 to 60 kpc and the cut radius is fixed to 1000 kpc. Its ellipticity can take any value from 0.3 to 0.6, its orientation can vary from 105° to 115° and the velocity dispersion can vary from 400 to 1000 km s^{-1}.

Table 1. Further details for *Ares* and *Hera* clusters.

Cluster name	z	Cluster galaxies	Images	Sources
Ares	0.5	330	242	85
Hera	0.507	337	65	19

4.1.2 *PIEMD–PIEMD*

This reconstruction is the same as the previous one but we add two BCG components (modelled using the PIEMD profile) whose coordinates are fixed to main large-scale haloes'. These BCGs are modelled separately and do not follow the scaling relations (Newman et al. 2013).

The BCG centred in the first clump (δRA, δDec) = (+20.0 arcsec, −32.0 arcsec) can have a core radius between 3.5 and 50.0 kpc, the cut radius can vary from 25 to 320 kpc, the ellipticity from 0.2 to 0.7, the orientation can vary from 120° to 180° and the velocity dispersion can vary from 100 to 500 km s^{-1}.

As for the BCG centred at (δRA, δDec.) = (−40.0 arcsec, +40.0 arcsec): Its core radius can vary between 25 and 35 kpc, the cut radius can vary from 670 to 730 kpc. Its ellipticity is allowed to vary between 0.4 and 0.65 kpc, the orientation angle can vary from 105° to 115° and the velocity dispersion can vary from 80 to 420 km s^{-1}.

4.1.3 *NFW*

This reconstruction, as in Section 4.1.1, has two cluster-scale haloes but modelled using an NFW density profile instead.

The clump fixed at (δRA, δDec.) = (+20.0 arcsec, −32.0 arcsec) away from the cluster centre can have an orientation between 130° and 140°, its velocity dispersion can vary between 500.0 and 2000.0 km s^{-1} and its scale radius varies from 50 to 280 kpc, and we let the ellipticity vary from 0.2 to 0.6.

The ellipticity of the second clump – fixed at (δRA, δDec.) = (−40.0 arcsec, +40.0 arcsec) – can reach values up to 0.6, the orientation is set between 105 and 115°, the velocity dispersion is set to vary between 500 and 1800 km s^{-1} and its scale radius can vary between 50 and 280 kpc.

Even if the ellipticity intervals are quite large for both clumps, after optimization, these clumps do no reach values $e \gtrsim 0.4$. This is in agreement with Golse & Kneib (2002), who demonstrated that the NFW profile is ill-defined for large ellipticities as defined in LENSTOOL. This will remain true for both clusters NFW models.

4.1.4 *NFW–PIEMD*

We add two BCG components (modelled with a PIEMD profile) to the reconstruction in Section 4.1.3, in the same way as in Section 4.1.2. At this point, we realized that the NFW models were a better fit to ARES (than PIEMD for the large-scale haloes) as the improvement of the logarithm of the Evidence in Table 2 shows (confirmed as the profile used to generate *Ares*'s large-scale haloes upon the unblinding during the challenge). Therefore, the following models, more complex as they also take into account more distant structures or an Hernquist profile, were modelled using only the NFW profile for the large-scale haloes.

Table 2. Mean value of the logarithm of the likelihood and image plane rms (arcsec), reduced χ^2, the logarithm of the Evidence and the BIC (Bayesian information criterion) for different *Ares* strong lensing models. In the 'Model' column, we specify the density profile chosen for the DM haloes and BCGs, respectively. Finally, we also show the best-fitting values of the cosmological parameters Ω_M and w. More details of these models are given in Section 4.1.

Model	log(Likelihood)	rms (arcsec)	Reduced χ^2	log(Evidence)	BIC	AICc	Ω_M	w
PIEMD–no BCGs	76.61	1.80	1.27	33.51	−15.23	−24.44	$0.24^{+0.04}_{-0.03}$	$-1.57^{+0.15}_{-0.23}$
PIEMD–PIEMD	117.00	0.66	0.81	61.46	−32.77	−37.93	$0.28^{+0.10}_{-0.06}$	$-1.02^{+0.05}_{-0.02}$
NFW–no BCGs	144.83	0.60	0.61	95.78	−163.17	−97.35	$0.31^{+0.06}_{-0.06}$	$-1.09^{+0.10}_{-0.24}$
NFW–PIEMD	155.28	0.73	0.58	100.70	−109.33	−83.76	$0.24^{+0.05}_{-0.05}$	$-1.03^{+0.06}_{-0.14}$
NFW–PIEMD + SUBS	166.53	0.55	0.48	107.65	−45.59	−47.14	$0.19^{+0.06}_{-0.03}$	$-0.99^{+0.05}_{-0.10}$
NFW–HERNQUIST	145.54	0.58	0.61	93.58	−118.59	−78.97	$0.27^{+0.03}_{-0.11}$	$-1.05^{+0.06}_{-0.27}$
NFW–HERNQUIST + SUBS	168.31	0.64	0.50	111.43	−60.65	−55.69	$0.33^{+0.03}_{-0.04}$	$-1.25^{+0.13}_{-0.12}$
NFW–PIEMD + shapes	135.56	0.57	0.65	79.82	−87.14	−64.04	$0.22^{+0.10}_{-0.05}$	$-0.90^{+0.07}_{-0.12}$

The last two models are 'post-unblinding' models. Once the input convergence map was made available, we discovered some structures in the outskirts of the cluster, out of the *HST* field of view (see Fig. 1).

4.1.5 NFW–PIEMD + SUBS

This model, together with the model 4.1.7, includes six additional DM components located in the outskirts of the cluster. This model is exactly like Section 4.1.4 with the additional substructures.

The substructures are located at (δRA, δDec.) = (+110.9 arcsec, −118.4 arcsec); (+84.3 arcsec, −138,8 arcsec); (+136.0 arcsec, −127,98 arcsec); (+57.6 arcsec, −83.2 arcsec); (−108,9 arcsec, +136,9 arcsec); and (+111.3 arcsec, −75.0 arcsec) away from the cluster's centre. They are modelled with NFW profiles, considered as spherical, their velocity dispersions are allowed to vary between 100 and 500 km s^{-1} and their scale radius varies between 20 and 200 kpc.

4.1.6 NFW–HERNQUIST

The large-scale clumps are modelled as in Section 4.1.4. The BCGs, on the other hand, have a different density profile: the Hernquist profile, as described in Section 3.1.

The BCG located at (δRA, δDec.) = (+20.0 arcsec, −32.0 arcsec) has a core radius between 1.0 and 40.0 kpc. Its ellipticity can reach values up to 0.4, the orientation varies between 90° and 180° and the velocity dispersion can vary from 100 to 400 km s^{-1}.

For the BCG located at (δRA, δDec.) = (−40.0 arcsec, +40.0 arcsec), the core radius is allowed to vary between 1.0 and 40.0 kpc, the ellipticity can vary from 0.1 to 0.7, the orientation can vary from 90° to 180° and the velocity dispersion can vary between 90 and 400 km s^{-1}.

4.1.7 NFW–HERNQUIST +SUBS

For this model, the large-scale haloes and BCGs components are modelled as in Section 4.1.6 and the distant substructures are modelled as in Section 4.1.5.

4.1.8 NFW–PIEMD + shapes

Last, we consider how taking into account the shapes of the galaxy-scale haloes impacts the reconstruction. The large-scale haloes and BCGs components are modelled as in Section 4.1.4 but we use SExtractor (Bertin & Arnouts 1996) to measure the semimajor and semiminor axis of the fitted ellipsis of each cluster galaxy, which are then taken into account in the modelling.

4.2 *Hera*

In spite of the unimodal appearance of the cluster and its convergence contours in the bottom panel of Fig. 1, Hera is also modelled as a bimodal cluster. Indeed, if this cluster is fitted with only one DM central clump, the χ^2 is twice larger.

The cluster members of the input catalogue were taken into account with $m_{F814W} < 24.0$ and were modelled in the same way for all models: parametrized by a PIEMD profile with a fixed core radius of 0.15 kpc, a velocity dispersion $\sigma^\star$ allowed to vary between 60.0 and 100.0 km s^{-1}, a cut radius $r^\star_{cut}$ with values from 1.0 to 200.00 kpc, considered spherical, with $mag0 = 19.8$, and following the scaling relations (Faber & Jackson 1976).

Moreover, in the same way as for *Ares*, four massive galaxies were more carefully modelled (indicated as green diamonds in Fig. 1), using a PIEMD density profile and fitted in the same way for each model.

The galaxy labelled as 1 in Fig. 1 is located at (δRA, δDec.) = (+25.91 arcsec, +26.98 arcsec) away from the cluster centre, considered spherical, with an angle position of 72° and a core radius of 0.07 kpc. Its cut radius r_{cut} can vary from 10.0 to 100.0 kpc and its velocity dispersion σ can vary from 50.0 to 200.0 km s^{-1}.

The next one, labelled as 2 in Fig. 1, is positioned at (δRA, δDec.) = (+60.38 arcsec, +20.2 arcsec), with a core radius of 0.31 kpc. Its ellipticity can go up to 0.5, the angle position is set between −65.0 and 90.0, the cut radius r_{cut} can vary from 30.0 to 400.0 kpc and its velocity dispersion σ can vary from 50.0 to 500.0 km s^{-1}.

The third halo of a massive galaxy (labelled as 3 in Fig. 1) is located at (δRA, δDec.) = (−19.665 arcsec, −11.078 arcsec) with a core radius of 0.1 kpc. Its ellipticity can vary between 0.1 and 0.7, the angle position can vary from 22° to 90°, the cut radius r_{cut} is allowed to vary between 8.0 and 150 kpc and the velocity dispersion σ can vary from 400 to 200.0 km s^{-1}.

Finally, the last one is a very massive galactic halo (labelled as 4 in Fig. 1) having a significant impact in the total mass distribution profile. This halo is located at (δRA, δDec.) = (−7.0 arcsec, −46.0 arcsec) with a core radius of 0.1, an ellipticity set between 0.1 and 0.7, an angle position allowed to vary between 75.0 to 90.0, a cut radius r_{cut} between 8.0 and 250.0 kpc and a velocity dispersion σ between 40.0 and 600.0 km s^{-1}.

All the multiple images provided were included in the model with a positional uncertainty of 0.5 arcsec and the optimization is

performed in the source plane. As for *Ares*, we checked that image plane and source plane models were giving similar results.

4.2.1 PIEMD

This model contains two large-scale haloes whose coordinates are placed at (δRA, δDec.) = (0.00 arcsec, 0.00 arcsec) and (δRA, δDec.) = (+20.0 arcsec, +2.00 arcsec) away from the centre of the cluster, respectively (and no BCGs). They are allowed to move around by ±2.00 arcsec.

For the first clump, we let the core radius vary from 5 to 37 kpc with a fixed cut radius of 3000 kpc. The ellipticity of this halo is allowed to vary values from 0.0 and 0.7, its orientation angle can vary from 0.0 to 180.0 and its velocity dispersion can vary from 400 to 1500 km s^{-1}. The second DM clump can have a core radius with any value from 5 to 40 kpc and a cut radius, also fixed, of 3000 kpc. Its ellipticity can take any value from 0.0 to 0.7, its orientation can vary from 0.0 to 180° and the velocity dispersion can vary from 600 to 1000 km s^{-1}.

4.2.2 PIEMD–PIEMD

We add two BCG components (modelled with a PIEMD profile) to the reconstruction in Section 4.2.1.

The BCG fixed at (δRA, δDec.) = (0.0 arcsec, 0.0 arcsec) has a fixed core radius of 0.5 kpc and a cut radius that can vary from 30.0 to 450.0 kpc. Its ellipticity can go from 0.0 to 0.4, the orientation can vary from 5° to 90° and the velocity dispersion can vary from 100 to 500 km s^{-1}.

The (δRA, δDec.) = (+20.0 arcsec, +2.00 arcsec) centred BCG has a fixed core radius of 1.0 kpc, the cut radius can vary from 20.0 to 400.0 kpc, the ellipticity can vary from 0.0 to 0.3, the orientation angle can vary from 12.0 to 90.0 and the velocity dispersion can vary between 100 and 500 km s^{-1}.

4.2.3 PIEMD–HERNQUIST

In this reconstruction, the two added BCGs to the model 4.2.1 are modelled using an Hernquist density profile.

The BCG located at (δRA, δDec.) = (0.0 arcsec, 0.0 arcsec) has a fixed core radius of 0.5kpc. Its ellipticity can reach values up to 0.4, the orientation angle is allowed to vary from 10° to 90° and the velocity dispersion can vary from 100 to 500 km s^{-1}.

For the BCG located at (δRA, δDec.) = (+20.0 arcsec, +2.00 arcsec), the core radius is fixed to 1.0 kpc, the ellipticity can vary from 0.0 to 0.3, the orientation angle can vary from 6° to 90° and the velocity dispersion can vary between 100 and 500 km s^{-1}.

4.2.4 NFW

This model, as in Section 4.2.1, has two large-scale haloes but modelled using an NFW density profile instead.

For the clump located at (δRA, δDec.) = (0.0 arcsec, 0.0 arcsec), aligned with the centre of the cluster, we let the ellipticity vary from 0.0 to 0.7 and its orientation can vary from 50° to 180°. It can have a velocity dispersion between 500 and 1500 km s^{-1} and scale radius varying from 50 to 280 kpc.

The ellipticity of the second clump [located at (δRA, δDec.) = (+20.0 arcsec, +2.00 arcsec)] can reach a value from 0.1 up to 0.7, the orientation is set between 0.0 and 180°, the concentration is set between 50 and 1500 km s^{-1} and its scale radius is set from 50 to 280 kpc.

4.2.5 NFW–PIEMD

For this model, the large-scale haloes are modelled as in Secion 4.2.4 and BCGs components are modelled as in Section 4.2.2

4.2.6 NFW–HERNQUIST

For this last model, the large-scale haloes are modelled as in Section 4.2.4 and BCGs components are modelled as in Section 4.2.3.

5 ESTIMATORS

In this section, we present four estimators used to assess the quality of our reconstructions and quantify the impact of the systematic errors arising from the choice of density profiles and configurations in the modelling of both *Ares* and *Hera*.

5.1 Statistical quality assessment

We use several statistical estimators to assess the quality of each reconstruction: the logarithm of the Likelihood, the rms (in arcseconds) in the image plane, the reduced χ^2, the logarithm of the Evidence and the BIC (following Lagattuta et al. 2017). The results are summarized in Tables 2 and 3 for *Ares* and *Hera*, respectively.

We would like to draw the reader's attention to the fact that all estimators are consistent with each other but using the logarithm of the Evidence appears to be a better way to differentiate between models. On the other hand, the rms does not necessarily reflect the improvement of the reconstruction. However, both the BIC and AICc strongly penalize models including the distant substructures in the modelling. Indeed, the greater number of free parameters in these models would not be justified even though they improve the mass distribution at large radii as well as tighten the cosmological constraints (see Figs 2 and 6, respectively).

As expected, in the case of *Ares*, considering an NFW profile for the large-scale potentials improves the modelling and Model 7 (NFW–HERNQUIST + subs) gives the best fit for this cluster as it is the closest to the true mass distribution. We show that taking into account the substructures in the cluster's outskirts for *Ares* leads to an improvement in the modelling of ∼20 per cent. Wide field imaging around strong lenses is thus crucial in order to detect such structures and include them in future reconstructions.

For the model NFW–PIEMD + shapes, we see that this additional information worsens the modelling of *Ares* with a lower value of log(Evidence) than the model without. This is in agreement with the true model as input cluster galaxies are spherical (Meneghetti et al. 2016), thus showing that our modelling technique is sensitive to the shapes of cluster members. As for *Hera*, the resulting fit is similar for all models. As expected, the reduced χ^2 values are larger than those for *Ares*, as the latter was modelled parametrically (assuming that light traces mass), thus better suited for our modelling technique.

5.2 Density profiles

In this section, we present the comparison between the radial density profile for all of our models and the true one.

In order to compare our projected mass maps to the input convergence map (with sources assumed at $z_s = 9$), the latter is normalized by multiplying the value of every pixel by $\Sigma_{\rm crit}$ (at $z_s = 9$, and assuming the input cosmology) in order to have the associated surface mass density $\Sigma(x, y)$ as

$$\Sigma(\boldsymbol{\theta}) = \kappa(\boldsymbol{\theta})\Sigma_{\rm crit}, \tag{9}$$

Table 3. Same as Table 2 for the *Hera* cluster. More details of these models are given in Section 4.2.

Model	log(Likelihood)	rms (arcsec)	Reduced χ^2	log(Evidence)	BIC	AICc	Ω_M	w
PIEMD–no BCGs	−35.93	0.99	2.85	−77.20	207.51	126.42	$0.54^{+0.14}_{-0.20}$	$-1.15^{+0.18}_{-0.73}$
PIEMD–PIEMD	−19.44	1.21	2.64	−64.97	210.71	151.36	$0.41^{+0.12}_{-0.09}$	$-1.55^{+0.58}_{-0.09}$
PIEMD–HERNQUIST	−31.80	0.96	2.98	−71.60	226.38	152.24	$0.44^{+0.16}_{-0.20}$	$-1.05^{+0.15}_{-0.57}$
NFW–no BCGs	−43.44	0.98	2.71	−79.81	222.53	133.93	$0.71^{+0.09}_{-0.30}$	$-1.30^{+0.64}_{-0.37}$
NFW–PIEMD	−46.47	1.00	3.03	−85.40	264.77	178.39	$0.33^{+0.18}_{-0.10}$	$-1.19^{+0.21}_{-0.50}$
NFW–HERNQUIST	−44.59	0.99	2.96	−81.40	251.96	165.02	$0.32^{+0.22}_{-0.07}$	$-1.48^{+0.47}_{-0.24}$

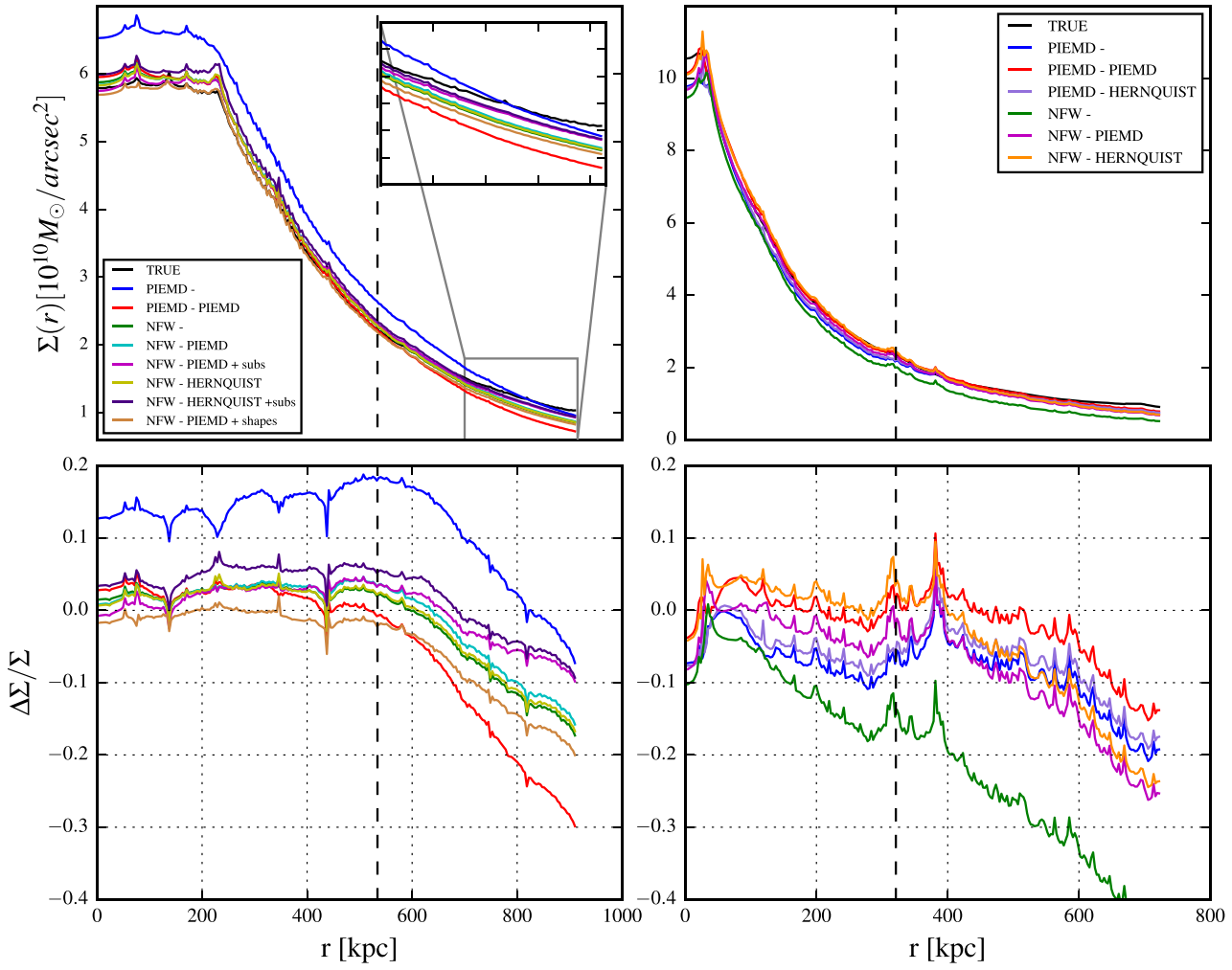

Figure 2. Top panels: density profiles for each model of Section 4.1 for *Ares* (left-hand panel) and of Section 4.2 for *Hera* (right-hand panel). Bottom panels: relative error of each model on the density for *Ares* (left-hand panel) and *Hera* (right-hand panel). The vertical dashed lines represent the radius below which there are multiple images.

where

$$\Sigma_{\rm crit} = \frac{c^2}{4\pi G}\frac{D_{\rm s}}{D_{\rm l}D_{\rm ls}}. \quad (10)$$

with c being the speed of light, G is the Newtonian constant and $D_{\rm s}$, $D_{\rm l}$ and $D_{\rm ls}$ are the angular diameter distances between observer–source, observer–lens and source–lens, respectively.

Our maps and the true one have the same field of view and spatial resolution.

For each model of Section 4, we compute the radial density profile and compare it to the true profile in Fig. 2.

(i) *Ares*: Regardless of the profile used for the large-scale potentials and BCGs, the mass distribution is well constrained within ∼5 per cent inside the radius below which there are constraints – defined as the radius of the circle enclosing all multiple images (except for the model PIEMD, which is overestimating the cluster's density by ∼14 per cent). However, the density profile in the outskirts of the *Ares* cluster (beyond the black vertical dashed line, representing

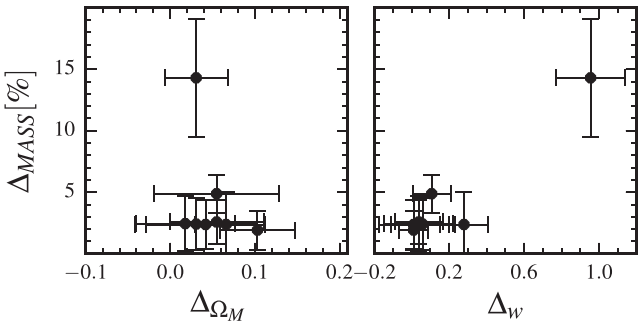

Figure 3. Bias on the total mass inside the radius below which there are multiple images for each model of Section 4.1 for *Ares* as a function of the bias on the estimations of the cosmological parameters Ω_M (left-hand panel) and w (right-hand panel).

the radius containing multiple images) tends to be underestimated by up to ∼30 per cent. Including the substructures in the modelling helps to better constrain the mass distribution, with an improvement between ∼5 per cent and ∼20 per cent in the most distant regions of the cluster. On average, the statistical error is about 5 per cent of the signal for all models, although for Model 1 (PIEMD), it varies from 4 per cent in the inner regions to up to 8 per cent in the outskirts.

We also show in Fig. 3 the bias on the estimation of Ω_M and w (i.e. the absolute error between the true value and the mode of the distribution for each of our models; see Fig. 6) depending on the bias on the total mass (inside the radius below which there are multiple images) for all of our models of Section 4.1 for *Ares* (as the constraints on $\Omega_M - w$ for *Hera* in Fig. 7 are too wide to draw any conclusion). As expected, the larger the bias on the mass is, the larger the bias on the cosmological parameters. The bias on the mass is therefore a cosmological quality estimator. Also, we show that the equation-of-state parameter w is more affected by a larger bias on the cluster mass than Ω_M. This result is in agreement with Golse et al. (2002), where the authors showed that the matter density parameter Ω_M can be better constrained than Ω_Λ for a set of simple mock strong lenses. In this work, we confirm that Ω_M is less sensitive to the modelling than w, even for more complex clusters such as *Ares*.

(ii) *Hera*: As expected, *Hera* is less constrained inside the radius below which there are multiple images (within ∼5–15 per cent) and the density profile at the very core of the cluster differs significantly from the true distribution. The density profile is well constrained by all models up to the radius containing multiple images. In the outskirts of the cluster, the density profile differs from the true distribution from ∼10 per cent up to ∼40 per cent depending on the model. The statistical error is larger than for *Ares*, especially for models 3, 4 and 6, for which it varies from ∼5 per cent in the centre up to 10–19 per cent in the outskirts.

5.3 Relative bias on magnification

In this section, we detail how we compute the magnifications from our lens models and compare them to the true values.

If we consider sources smaller than the angular scale on which the lens properties change, the Jacobian matrix describing the distortion (in shape and size) of images is then

$$\mathcal{A}(\theta) = \begin{pmatrix} 1-\kappa-\gamma_1 & \gamma_2 \\ \gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix}, \tag{11}$$

where κ is the convergence and γ_1, γ_2 are the first and second shear components, respectively. The magnification is then the inverse of the determinant of $\mathcal{A}$:

$$\mu = (det\mathcal{A})^{-1} = \frac{1}{|(1-\kappa)^2 - |\gamma|^2|}, \tag{12}$$

where *mu* is the magnification of the source.

(i) True magnification: To obtain the true magnification for each multiple image, we use the true convergence and shear maps at $z_s = 9$ at any location in the image plane covering a field of view of 300×300 arcsec^2 for *Ares* and 227.16×227.16 arcsec^2 for *Hera*.

However, to have the true magnification at the source's redshift, we multiply the true convergence κ and shear components γ_1, γ_2 by a normalizing factor:

$$\frac{D_{OS}}{D_{LS}}(z_s = 9) \times \frac{D_{LS}}{D_{OS}}(z_s = z). \tag{13}$$

Finally, we interpolate this map to get the magnification at the position of all input multiple images.

(ii) Measured magnification: The magnification for each multiple image is measured for all of our reconstructions with the LENSTOOL software using equations (11) and (12).

We are interested in how the measured magnification evolves with an increasing number of multiple images. To do so, we run SExtractor (Bertin & Arnouts 1996) to extract and measure the magnitudes on the *F814W* image (deeper, thus more arcs detected) of all the multiple images that are then divided into five different catalogues with increasing magnitude thresholds in the *F814W* filter (being uniformly distributed in the field).

We then run each lens model presented in Sections 4.1 and 4.2 with all of the new multiple image catalogues (with different magnitude thresholds). We compute the relative bias on magnification per multiple image as well as the 25th and 75th percentiles of this distribution. Finally, for each magnitude catalogue, we compute an average bias on the magnification with

$$\delta\mu = p75\left(\frac{\mu^{fit} - \mu^{true}}{\mu^{true}}\right) - p25\left(\frac{\mu^{fit} - \mu^{true}}{\mu^{true}}\right), \tag{14}$$

where μ^{fit} is the mode of the magnification distribution per catalogue.

We show in Fig. 4 the relative bias on magnification as a function of the number of images taken into account in the modelling for both clusters (top panel) and the precision of our measurements (bottom panel). The bias on the magnification is of the order of ∼20 per cent for *Ares*. This bias can go up to 40 per cent for the smallest bins. The magnifications of the multiple images in Hera are less constrained with an average bias of ∼30 per cent. The relative bias on magnification is not reduced by an increasing number of constraints for either of the clusters. However, the precision on magnification improves by a factor of ∼4 when increasing the constraints in the modelling as stated in Jauzac et al. (2014). All models provide similar measurements of the magnification and modelling galaxy clusters in a more complex way is not translated into an improvement in the accuracy or precision of magnification.

Taking into account the substructures in the cluster's outskirts leads to a very localized improvement of the magnification bias of ∼15 per cent but remains constant within the cluster's core.

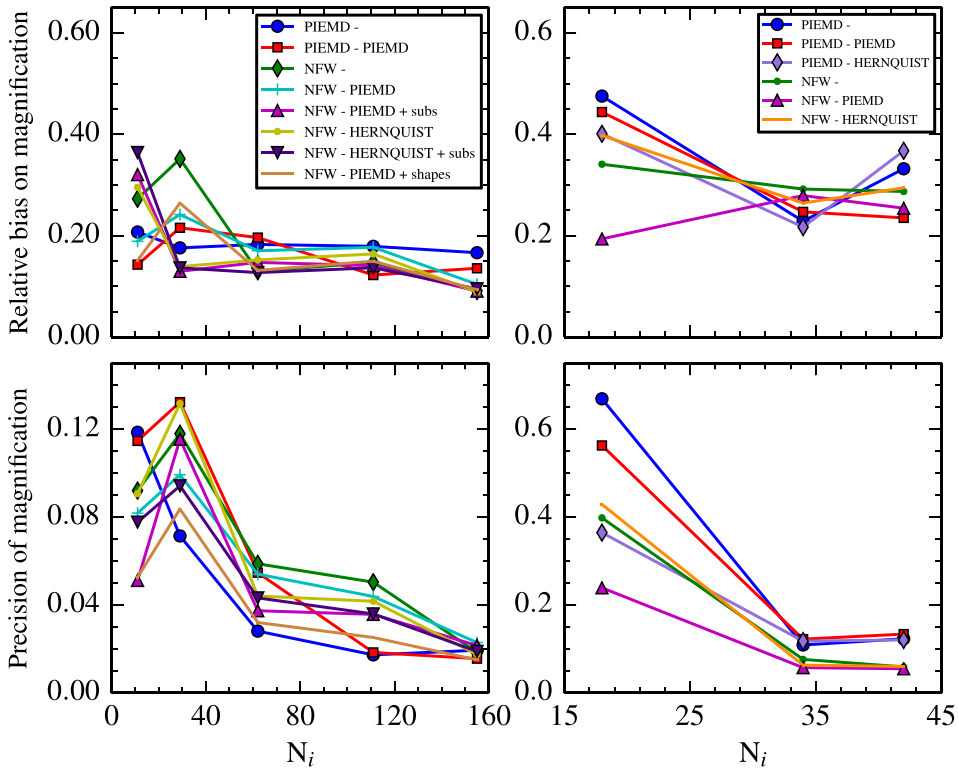

Figure 4. Top panels: relative error on magnification as a function of the increasing number of multiple images taken into account in the modelling. Left-hand panel: *Ares*. Right-hand panel: *Hera*. Bottom panels: relative precision on the magnification measurements as a function of the increasing number of multiple images. Left-hand panel: *Ares*. Right-hand panel: *Hera*.

5.4 Cluster ellipticity and orientation

We use the mass maps generated for each model and the normalized true convergence map (as explained in Section 5.2) to compute the cluster ellipticity as a function of the major semi-axis and its position angle for each model of Section 4.

We set a list of iso-surface density thresholds to fit an ellipse to each contour. As both clusters are bimodal, the highest iso-density contour is set to be the one that encloses both clumps.

The ellipticity is then computed as follows:

$$e = \frac{a^2 - b^2}{a^2 + b^2}, \tag{15}$$

where a and b are the measured major and minor semi-axes, respectively. The radial profiles of the ellipticity and of the orientation angle for both clusters are shown in Fig. 5.

For each panel, the true values are given by the black line. Overall, the values of all the models are in good agreement with the true value in the inner regions (before the dashed vertical line, representing the radius up to where we have multiple images) but tend to differ in the outskirts as our models do not have constraints anymore, thus extrapolating from the core region. The ellipticity in the outskirts of *Ares* is underestimated but overestimated for *Hera*. The values of the cluster's orientation angle are recovered within 5° for *Ares* and within 10° for *Hera*.

This estimator does not allow us to discriminate between the different models as all models are in perfect agreement.

6 COSMOGRAPHY

In this section, we investigate the impact of considering different cluster modellings on the estimation of cosmological parameters.

We assess first if the choice of different density profiles for the cluster components (large-scale potentials and BCGs) has a significant impact on constraining the Ω_M and w parameter space.

We also analyse the systematic error introduced when using different redshift catalogues for the background sources (spectroscopic only, spectroscopy and photometry) for the lens modelling.

6.1 Estimation of Ω_M and w

As mentioned before, for each model, we have (taking into account all multiple images with spectroscopic redshifts), the cosmological parameters Ω_M and w are left as free parameters. The constraints obtained are shown in Fig. 6 for *Ares* and Fig. 7 for *Hera*.

Using either different density profiles for the large-scale clumps and BCGs provides similar constraints on the Ω_M and w parameter space (within the 1σ contours) if a large number of multiple images is available (with spectroscopic redshifts and with a positional error of 0.5 arcsec) and if the model is realistic enough (all but model 1:PIEMD, which has the largest mass bias in Fig. 3).

However, including in the modelling massive substructures in the cluster's outskirts translates into a decrease of the statistical errors as the cosmological contours are smaller: at the risk of biasing the results. Our work is in agreement with these previous studies showing that massive structures in the outskirts of clusters do

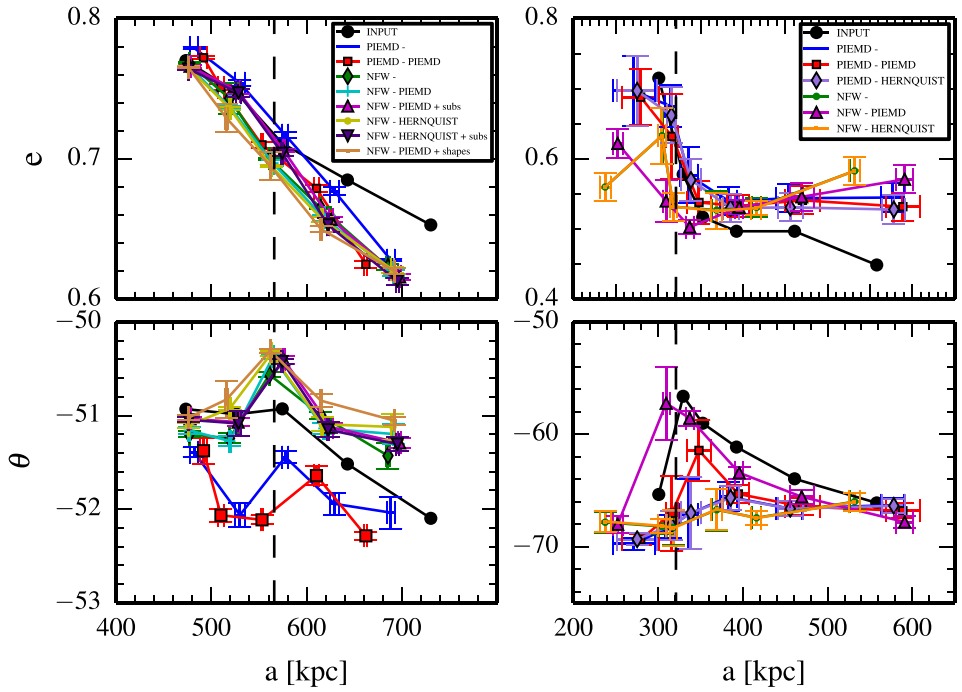

Figure 5. Top panels: radial profile of the cluster's ellipticity. Bottom panels: radial profile of the orientation angle. The dashed vertical lines represents the limit for which we have multiple images. Left panels: *Ares* cluster. Right panels: *Hera* cluster.

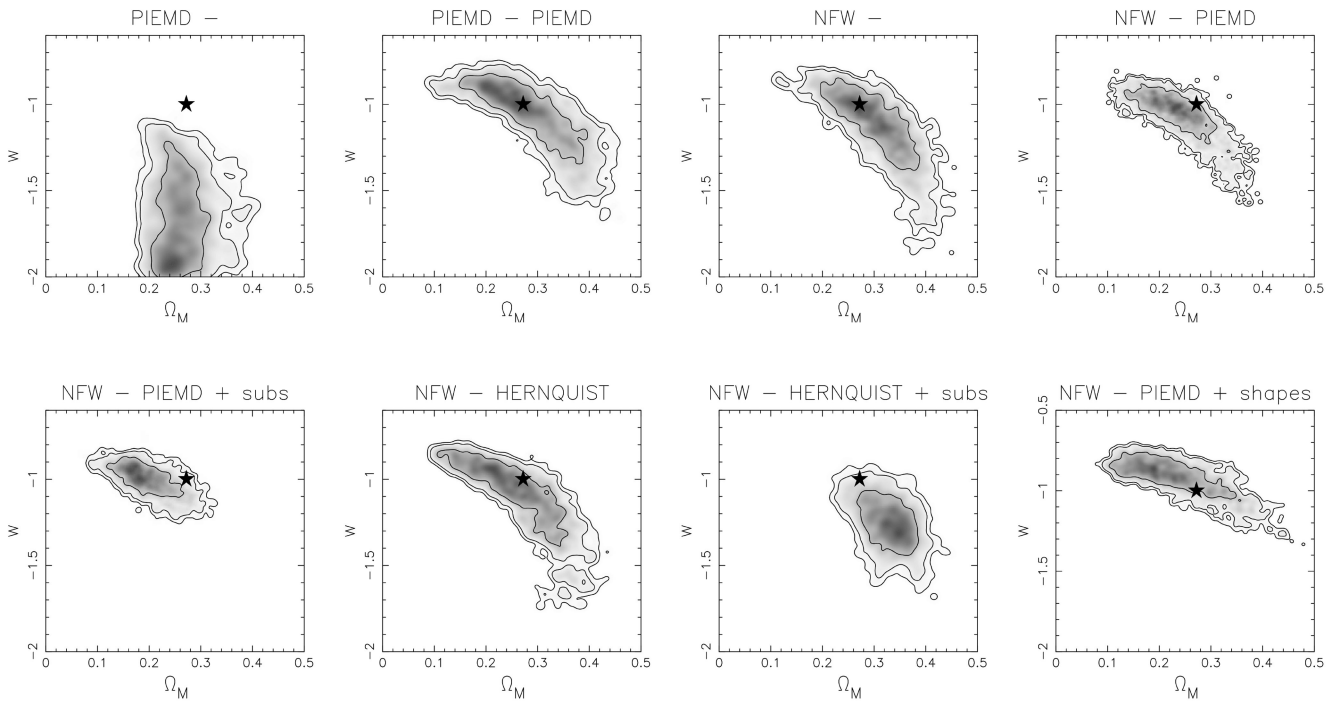

Figure 6. Estimation of cosmological parameters Ω_M–w for the *Ares* cluster using all multiple images. Above each panel, the density profile used to fit the smooth component and BCGs is specified. The plotted contours are the 1, 2 and 3σ confidence levels and the star indicates the true value.

impact not only the mass distribution but also the constraints on cosmological parameters yielding smaller contours. Not only do the line-of-sight structures introduce a systematic error in the strong lensing modelling (e.g. Bayliss et al. 2014; Giocoli et al. 2016) but also distant massive structures in the lens plane have a considerable impact in the position of multiple images (Tu et al. 2008; Limousin et al. 2010) and thus on the mass constraints. Indeed, in Mahler et al. (2017), the authors evaluate the impact of the

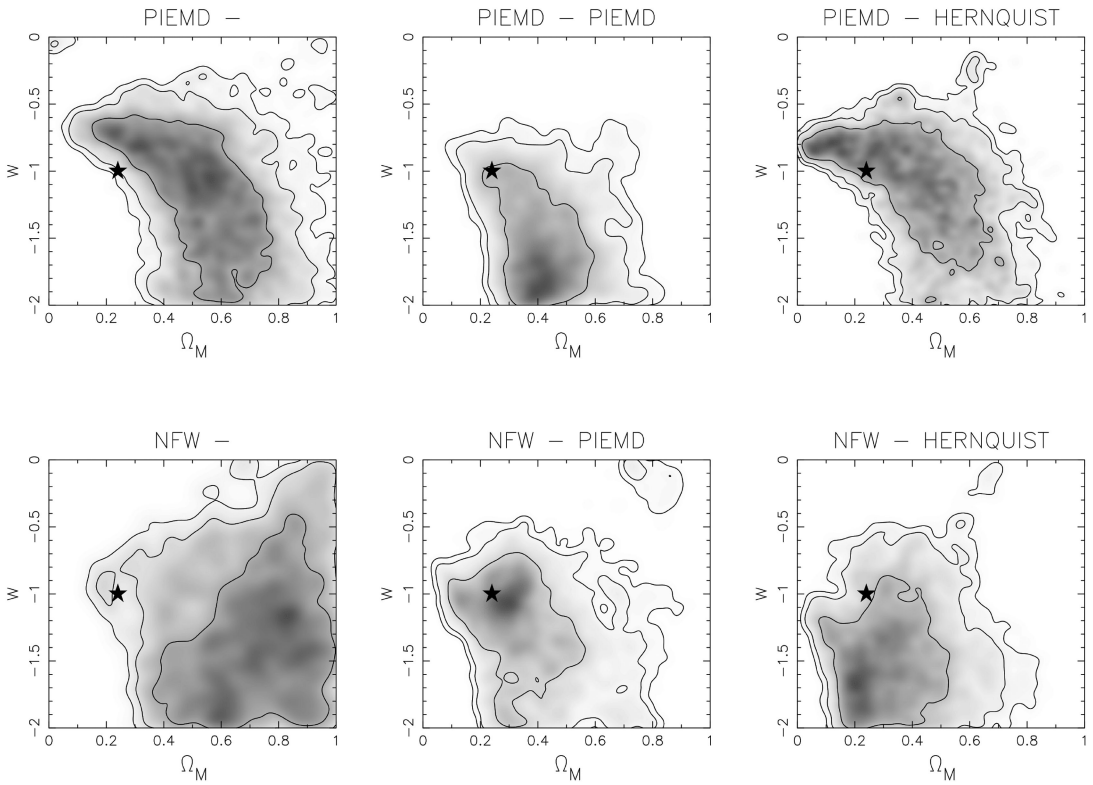

Figure 7. Estimation of cosmological parameters Ω_M–w for the *Hera* cluster using all multiple images. Above each panel, the density profile used to fit the smooth component and BCGs is specified. The plotted contours are the 1, 2 and 3σ confidence levels and the star indicates the true value.

presence of mass clumps in the outskirts of the cluster core of Abell 2744 (Jauzac et al. 2016) to change the mass profile by ∼6 per cent at 200 kpc. McCully et al. (2017) have shown that considering the environmental and line-of-sight perturbations should not be taken aside in the modelling as in doing so, the fit does not reproduce the input lens system parameters or the Hubble constant.

The contours obtained with *Hera* are, in general, much larger than those from *Ares*. This can be explained by the fact that the redshift range of *Ares*'s sources is twice larger than for *Hera*. By running one of *Ares* models with multiple images only up to $z \sim 3.5$, we checked that the contours were larger indeed. The widening of the contours might then not be due to the number of images taken into account as seen in Fig. 9. The left-hand panels show the constraints on the Ω_M–w parameter space for one of *Ares* configurations. Considering the same number of multiple images, but spanning a twice larger range of redshift, provides tighter constraints than those obtained with Hera.

Finally, we would like to point out that recovering cosmological parameters with strong lenses have been until now performed for unimodal clusters (or more simple clusters than the FF) that would be the preferred configuration for cosmography. Recently, Caminha et al. (2016) performed the first cosmography analysis with a Frontier Fields cluster, AS1063, which is the most relaxed cluster of the sample. By carefully selecting a sub-sample of secured multiple images, they achieve an rms of 0.3 arcsec and put constraints on the Ω_M, w and Ω_Λ parameters.

However, we show in this work that complex and multimodal clusters such as *Ares* can yield tight and competitive constraints on the Ω_M–w parameter space.

6.2 Redshift catalogues

In this section, we investigate how the estimation of robust cosmological parameters is affected by the availability of different sets of constraints. We extend this study for three of the models for *Ares* in Section 4.1 (models 2, 3, 6).

These analyses have been carried out only for the *Ares* cluster as it has three times more multiple images (see Table 1) and the range of redshift is twice wider than for *Hera*.

6.2.1 Redshift range

We investigate first whether there is a redshift range of background sources more efficient to recover the input cosmology and if photometric redshifts can complete our samples.

We split the redshift catalogue into four bins of redshift:

(i) *Bin1* as $z_s \leq 1.96$,
(ii) *Bin2* as $1.96 < z_s \leq 3.08$,
(iii) *Bin3* as $3.08 < z_s \leq 3.68$,
(iv) *Bin4* as $z_s > 3.68$,

and we build the new multiple images catalogues as follows: For each of the new four multiple images catalogues, we keep the images whose redshift is inside each bin as spectroscopic and the rest as photometric redshifts that have been virtually created with a precision of $0.04(1 + z)$ and assuming that there are no catastrophic errors. This precision is already achievable with the HFF data: Castellano et al. (2016) determined a typical photometric redshift error on the multiple images in Abell 2744 and MACS 0416 of 3–5 per cent. Of the 242 total images provided, we keep ∼60 with

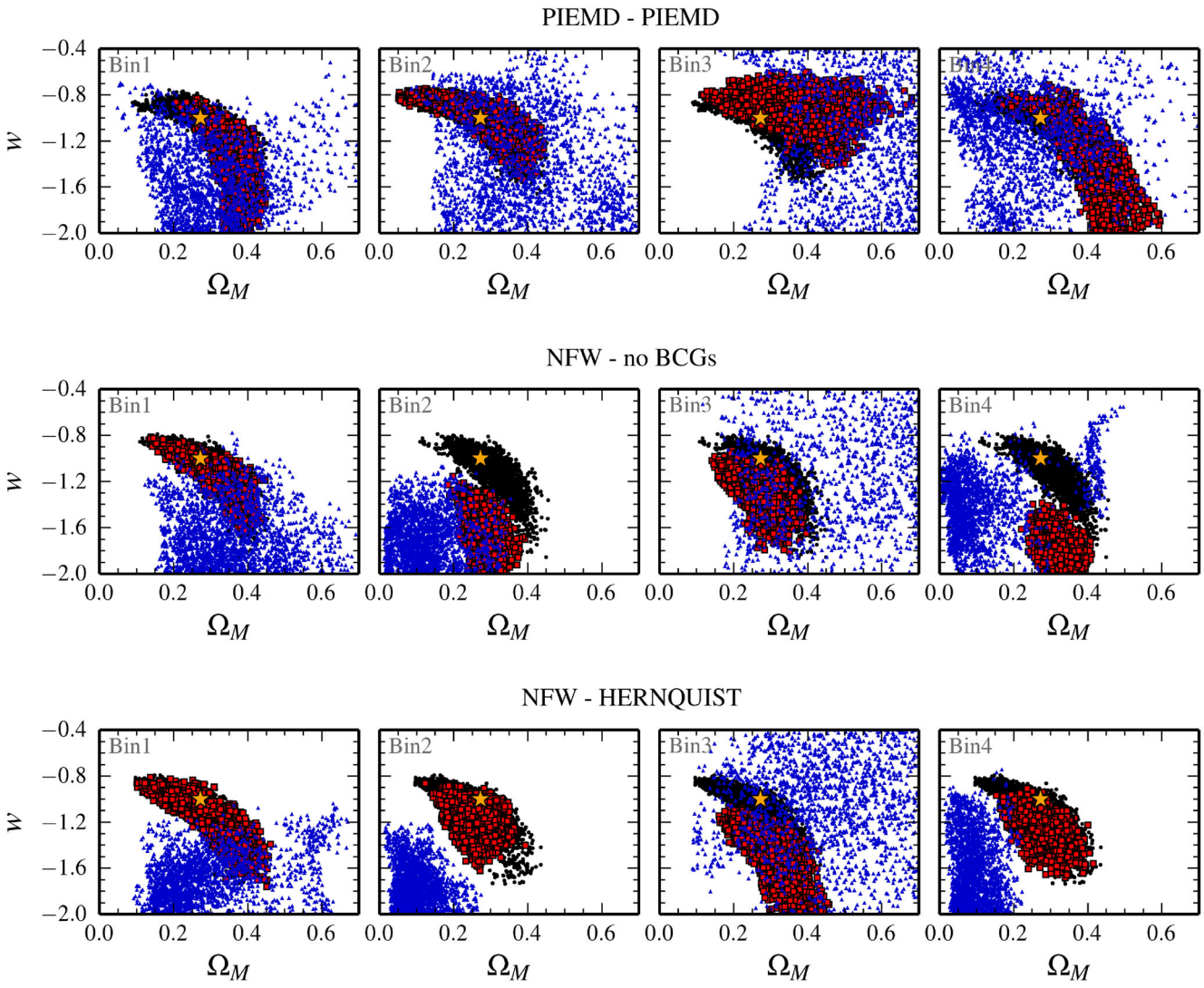

Figure 8. Constraints on the Ω_M and w parameters for *Ares*. The fiducial constraints (all multiple images with spectroscopic redshifts) are presented as filled black circles in Fig. 6. When considering only arcs within the redshift range specified in each panel with spectroscopic information (blue filled triangles), the constraints are obtained biased. If we consider not only arcs within the redshift range specified in each panel with spectroscopic information but also the remaining catalogues with photometric redshifts [with a $0.04(1 + z)$ precision], there is no bias on the cosmological constraints (shown by the filled red squares) or, at least, it is reduced.

spectroscopic information per bin and the rest with photometric accuracy.

We also study if it is preferable to use only spectroscopic information or add photometric redshifts as constraints so we also use multiple images catalogues with only images belonging to each bin (having then ∼60 images, with spectroscopy). We compare, for each bin of redshift stated above, the constraints on the Ω_M–w parameter space with photometric redshifts as additional information or not.

The results are shown in Fig. 8. For reference, the fiducial models of Fig. 6 are also shown (black constraints). First, we show that considering only spectroscopic redshifts from a certain redshift range (blue points) biases the estimation of cosmological parameters, whatever the redshift bin considered is, the bias being similar whichever the profile and/or configuration used are for the same bin considered.

By completing these spectroscopic redshifts catalogues with multiple images with photometric information (red points), spanning the ranges in redshift not covered by them, we recover unbiased cosmology (golden star) in most cases or at least reduce this bias. Note that if cosmological parameters are still biased, it is the equation-of-state parameter w that is more affected than Ω_M.

Photometric redshifts are then a useful piece of additional information to take into account in the modelling and completing the spectroscopic catalogue if the latter covers only a narrow redshift range.

6.2.2 *Photometric families*

Considering multiple images from a restricted range of redshift leads to an estimation of wider constraints (which are biased) on the Ω_M and w parameter space, whatever the mass distribution of the cluster is. This bias is due to the narrow range of redshifts considered and not due to a reduced number of multiple images considered as seen below.

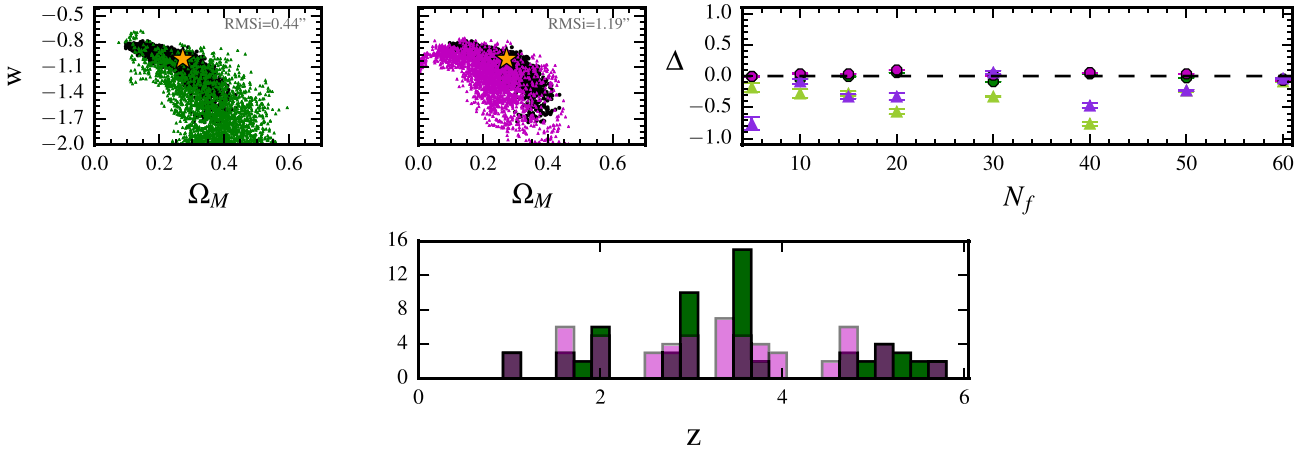

Figure 9. Top left-hand and middle panels: constraints on the Ω_M and w parameters for *Ares* for the model NFW–HERNQUIST. As filled black circles are plotted the fiducial constraints (all multiple images with spectroscopic redshifts); as green and magenta filled triangles are plotted the constraints from two models with a (slightly different) reduced catalogue of spectroscopic redshifts (shown in the bottom panel). Right-hand panel: systematic bias in the recovery of cosmological parameters as a function of increasing photometric families added in the modelling [with a $0.04(1+z)$ precision]. The filled triangles show the bias on the w parameter and the filled circles show the bias on the Ω_M parameter. The statistical error is represented by the error bars.

In this section, we consider as constraints a reduced catalogue of multiple images with spectroscopic redshifts by creating two similar catalogues (but with a different redshift distribution), which contain ∼60 images and spanning all the redshift range available (see histograms in Fig. 9). We have considered the same three mass models (id 2, 3 and 6) of Section 4.1 for *Ares*. For clarity, we show the results obtained for only one of them in Fig. 9 as the results were very similar for the three of them.

In the left-hand and middle panels, we show in black the constraints obtained with the fiducial model NFW–HERNQUIST of (Section 6); on the top are the constraints with the two reduced catalogues. As we can note, the cosmological parameters are unbiased.

We investigate if adding an increasing number of families of multiple images with photometric redshifts (spanning all the redshift range) in the modelling translates into an improvement in the cosmology estimation (i.e. a smaller systematic than statistical error). This is shown in the right-hand panels of Fig. 9 where the systematic and statistical errors (errors bars) are plotted as a function of the increasing number of photometric families taken in account. The coloured circles show the bias on Ω_M and the triangles show the bias on w. We observe the same behaviour for the three models, where Ω_M is less affected by w than the modelling and the statistical errors are smaller than the systematic uncertainties. w is systematically underestimated. We do not report, however, any trend with the increasing number of photometric families.

6.3 On the positional uncertainty of the multiple images

Throughout this paper, we have assumed an uncertainty of 0.5 arcsec for the position of multiple images, closer value to the rms in the image plane, thus providing a reduced $\chi^2 \sim 1$. We show in Fig. 10 the bias on the estimation of Ω_M and w for three positional errors of the multiple images assumed in the modelling for two models of Section 4.1 for *Ares*. This bias is the lowest for a positional uncertainty of the order of the rms. However, this figure also shows that underestimating the uncertainties on the observations can lead to biased constraints on the Ω_M–w parameter space.

7 CONCLUSIONS

We have analysed two mock galaxy clusters, *Ares* and *Hera* from the FF-SIMS Challenge (Meneghetti et al. 2016), both complex and bi-modal, comparable to the FF clusters. We have investigated the systematic errors in the strong lensing parametric modelling arising from the choice of the density profiles and configurations as well as from the availability of constraints (spectroscopic or photometric redshifts) and therefore the impact on the retrieval of robust cosmological parameters.

Our main conclusions are the following:

(i) Galaxy clusters are not isolated systems and can present large structures in the cluster outskirts (Jauzac et al. 2016; Foëx, Chon & Böhringer 2017). With this work, we provide further evidence that distant massive substructures in the lens plane of galaxy clusters have a significant impact on the mass distribution. Wide-field imaging around massive clusters is thus needed to account for these structures in the modelling. In an era of precise cosmology, we show that the cluster's environment cannot be ignored in order to yield a more precise mass reconstruction and therefore competitive constraints on Ω_M and w.

(ii) As expected, the smaller the bias on the mass, the smaller is the bias on cosmological parameters. The bias on the total mass is a quality indicator for the cosmological constraints. On the other hand, magnification, the cluster's ellipticity and orientation do not allow us to discriminate between models (assuming the same modelling technique).

(iii) Considering a positional error of 0.5 arcsec, the estimation of cosmological parameters is not affected by the choice of different mass profiles or configurations when a sufficient number of constraints is available ($n_{\rm im} > \sim 60$).

(iv) The bias on the estimation of cosmological parameters is the lowest for a positional uncertainty of the order of the rms. Underestimating the uncertainty on the observations can lead to biased constraints on the Ω_M–w parameter space.

(v) Considering multiple images, from a restricted range of redshift, leads to an estimation of biased cosmological parameters.

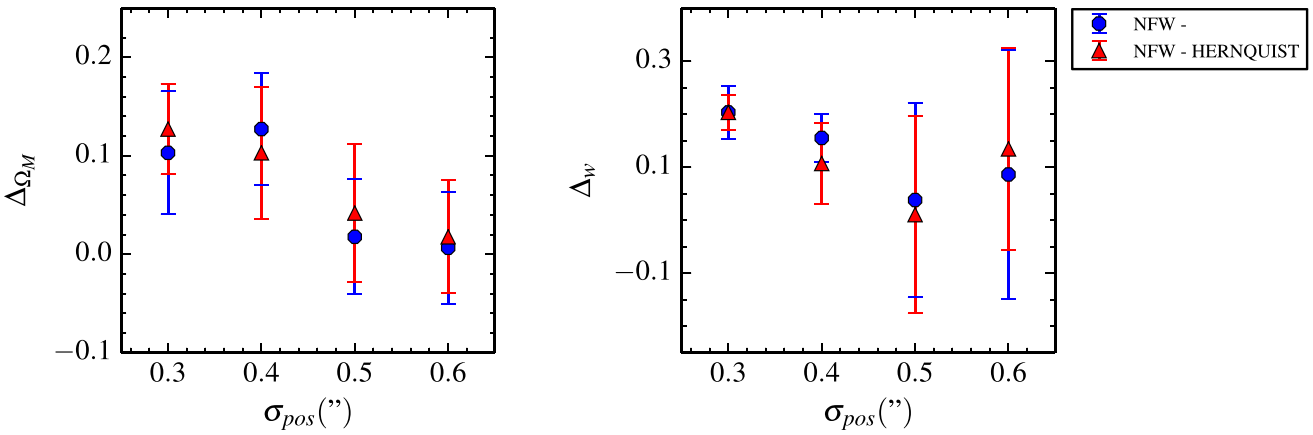

Figure 10. Bias on the estimation of Ω_M (left-hand panel) and w (right-hand panel) depending on the assumed positional error of the multiple images for two mass distributions for *Ares* (the NFW–HERNQUIST being more accurate)

Taking into account multiple images from a broader range of redshift with photometry information can correct this bias or, at least, reduce it.

(vi) We do not report any trend between an increasing number of photometric families taken into account in the modelling and a more precise estimation of Ω_M and w.

(vii) Ω_M is less sensitive to systematic errors than w, the latter being systematically underestimated when recovered biased.

Stronger constraints can be obtained by combining the estimates on Ω_M and w from several strong lensing clusters (D'Aloisio & Natarajan 2011). We show that not only unimodal (or simpler clusters than *Ares* and *Hera*; Jullo et al. 2010; Caminha et al. 2016) but also more complex and multimodal clusters can yield competitive constraints. Upcoming surveys such as *James Webb Space Telescope* will make strong lensing cosmography a very powerful tool by detecting an even larger number of arcs than currently available with *HST*.

ACKNOWLEDGEMENTS

We thank the referee for her/his useful comments that helped improving this work. AA is grateful to Gabriel Bartosch Caminha for useful and constructive discussions. We also thank M. Meneghetti, P. Natarajan and D. Coe for providing the *Ares* and *Hera* mock clusters for the Frontier Fields Lens Modelling Comparison Project.

This work has been carried out by to the support of the OCEVU Labex (ANR-11-LABX-0060) and the A*MIDEX project (ANR-11-IDEX-0001-02) funded by the 'Investissements d'Avenir' French government programme managed by the ANR.

This work was granted access to the HPC resources of Aix-Marseille Université financed by the project Equip@Meso (ANR-10-EQPX-29-01) of the programme 'Investissements d'Avenir' supervised by the Agence Nationale pour la Recherche.

We are also grateful to CNES for financial support. ML thanks CNRS for financial support.

CG acknowledges support from the Italian Ministry for Education, University and Research (MIUR) through the SIR individual grant SIMCODE, project number RBSI14P4IH

MJ acknowledges support by the Science and Technology Facilities Council [grant number ST/L00075X/1].

REFERENCES

Atek H. et al., 2015, ApJ, 814, 69
Bayliss M. B., Johnson T., Gladders M. D., Sharon K., Oguri M., 2014, ApJ, 783, 41
Bayliss M. B., Sharon K., Johnson T., 2015, ApJ, 802, L9
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Beutler F. et al., 2011, MNRAS, 416, 3017
Bouwens R. J., Oesch P. A., Illingworth G. D., Ellis R. S., Stefanon M., 2016, Astrophys. J., preprint (arXiv:1610.00283)
Caminha G. B. et al., 2016, A&A, 587, A80
Castellano M. et al., 2016, A&A, 590, A31
Coe D., Benítez N., Sánchez S. F., Jee M., Bouwens R., Ford H., 2006, AJ, 132, 926
Coe D. et al., 2013, ApJ, 762, 32
D'Aloisio A., Natarajan P., 2011, MNRAS, 411, 1628
Dalal N., Hennawi J. F., Bode P., 2005, ApJ, 622, 99
de Haan T. et al., 2016, ApJ, 832, 95
De Lucia G., Blaizot J., 2007, MNRAS, 375, 2
Despali G., Giocoli C., Bonamigo M., Limousin M., Tormen G., 2017, MNRAS, 466, 181
Diego J. M., Broadhurst T., Wong J., Silk J., Lim J., Zheng W., Lam D., Ford H., 2016, MNRAS, 459, 3447
Faber S. M., Jackson R. E., 1976, ApJ, 204, 668
Foëx G., Chon G., Böhringer H., 2017, A&A, 601, A145
Gil-Marín H. et al., 2016, MNRAS, 460, 4210
Gilmore J., Natarajan P., 2009, MNRAS, 396, 354
Giocoli C., Meneghetti M., Bartelmann M., Moscardini L., Boldrin M., 2012, MNRAS, 421, 3343
Giocoli C., Meneghetti M., Metcalf R. B., Ettori S., Moscardini L., 2014, MNRAS, 440, 1899
Giocoli C., Bonamigo M., Limousin M., Meneghetti M., Moscardini L., Angulo R. E., Despali G., Jullo E., 2016, MNRAS, 462, 167
Golse G., Kneib J.-P., 2002, A&A, 390, 821
Golse G., Kneib J.-P., Soucail G., 2002, A&A, 387, 788
Harvey D., Kneib J. P., Jauzac M., 2016, MNRAS, 458, 660
Hernquist L., 1990, ApJ, 356, 359
Heymans C. et al., 2012, MNRAS, 427, 146
Hildebrandt H. et al., 2017, MNRAS, 465, 1454
Host O., 2012, MNRAS, 420, L18
Jauzac M. et al., 2014, MNRAS, 443, 1549
Jauzac M. et al., 2016, MNRAS, 463, 3876
Johnson T. L., Sharon K., 2016, ApJ, 832, 82
Jullo E., Kneib J.-P., Limousin M., Elíasdóttir Á., Marshall P. J., Verdugo T., 2007, New J. Phys., 9, 447

Jullo E., Natarajan P., Kneib J.-P., D'Aloisio A., Limousin M., Richard J., Schimd C., 2010, Science, 329, 924
Kassiola A., Kovner I., 1993, ApJ, 417, 450 (PIEMD)
Kneib J.-P., Ellis R. S., Smail I., Couch W. J., Sharples R. M., 1996, ApJ, 471, 643
Kneib J.-P., Ellis R. S., Santos M. R., Richard J., 2004, ApJ, 607, 697
Lagattuta D. J. et al., 2017, MNRAS, 469, 3946
Limousin M., Kneib J.-P., Natarajan P., 2005, MNRAS, 356, 309
Limousin M. et al., 2010, A&A, 524, A95
Limousin M. et al., 2016, A&A, 588, A99
Link R., Pierce M. J., 1998, ApJ, 502, 63
Lotz J. M. et al., 2017, ApJ, 837, 97
McCully C., Keeton C. R., Wong K. C., Zabludoff A. I., 2017, ApJ, 836, 141
Magaña J., Motta V., Cardenas V. H., Verdugo T., Jullo E., 2015, ApJ, 813, 69
Mahler G. et al., 2017, MNRAS, preprint (arXiv:1702.06962)
Massey R. et al., 2007, ApJS, 172, 239
Meneghetti M. et al., 2008, A&A, 482, 403
Meneghetti M., Fedeli C., Pace F., Gottlöber S., Yepes G., 2010, A&A, 519, A90
Meneghetti M. et al., 2016, MNRAS, preprint (arXiv:1606.04548)
Monna A. et al., 2017, MNRAS, 466, 4094
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Newman A. B., Treu T., Ellis R. S., Sand D. J., 2013, ApJ, 765, 25
Peel A., Lin C.-A., Lanusse F., Leonard A., Starck J.-L., Kilbinger M., 2017, American Astronomical Society Meeting Abstracts, 430.01
Planck Collaboration XIII, 2016a, A&A, 594, A13
Planck Collaboration XIV, 2016b, A&A, 594, A14
Planelles S., Borgani S., Fabjan D., Killedar M., Murante G., Granato G. L., Ragone-Figueroa C., Dolag K., 2014, MNRAS, 438, 195
Richard J., Stark D. P., Ellis R. S., George M. R., Egami E., Kneib J.-P., Smith G. P., 2008, ApJ, 685, 705
Richard J., Kneib J.-P., Limousin M., Edge A., Jullo E., 2010, MNRAS, 402, L44
Riess A. G. et al., 1998, AJ, 116, 1009
Schwarz G., 1978, Ann. Stat., 6, 461
Sereno M., Giocoli C., Ettori S., Moscardini L., 2015, MNRAS, 449, 2024
Soucail G., Kneib J.-P., Golse G., 2004, A&A, 417, L33
Suyu S. H. et al., 2017, MNRAS, 468, 2590
Tu H., Limousin M., Fort B., Shu C. G., Sygnet J. F., Jullo E., Kneib J. P., Richard J., 2008, MNRAS, 386, 1169
Wang L., Li C., Kauffmann G., De Lucia G., 2006, MNRAS, 371, 537
Zitrin A., Broadhurst T., 2016, ApJ, 833, 25
Zitrin A. et al., 2015, ApJ, 801, 44

APPENDIX A: SOURCE REDSHIFT DISTRIBUTION

We show in Fig. A1 the redshift distribution of background sources for the *Ares* and *Hera* clusters.

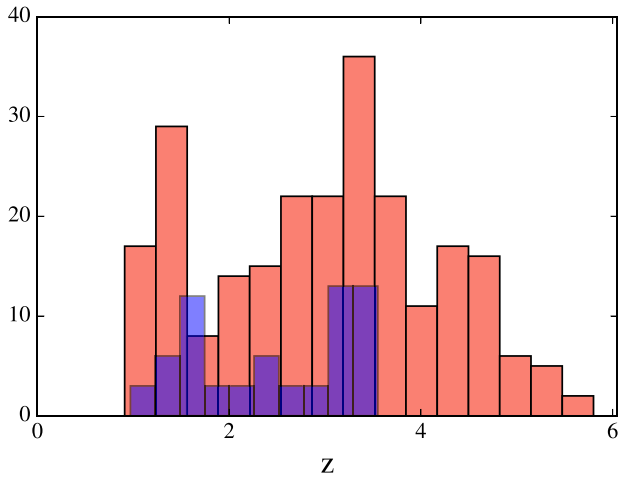

Figure A1. Resdhift distribution of the background sources for the strong lens *Ares* in pink and for *Hera* in blue.

This paper has been typeset from a TEX/LATEX file prepared by the author.